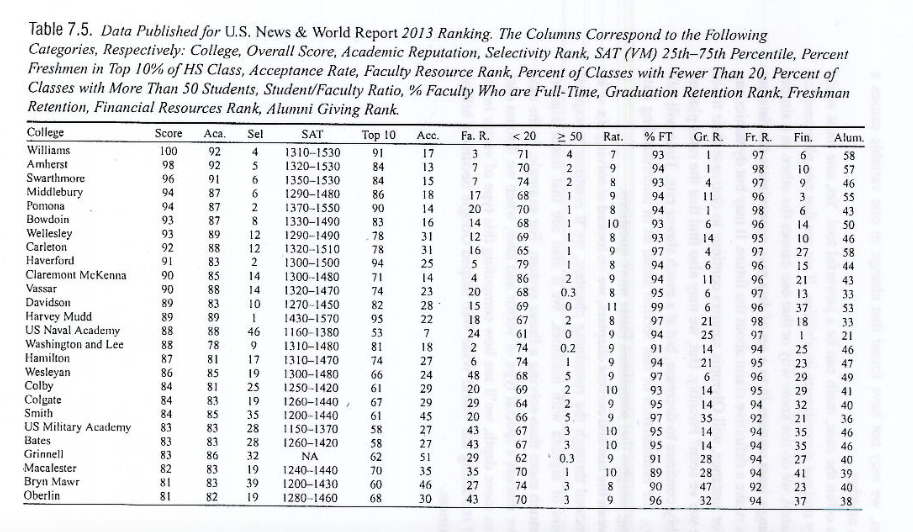
**College Rankings Project** Name:

You can open my college data file to find the following table from Life is Linear:



**Project Directions**

Create a Python program that does the following:

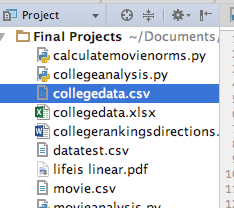
1. Uses *all* of the colleges in the above list except for Grinnell (who doesn’t publish SATs) and all of the variables above (Academic Reputation through Alumni Giving Rank) to create a matrix, A. In addition, create a vector, b, containing all of the colleges’ (except for Grinnell’s) US News World & Report Score (2nd column).

2. Solves for the weight of each ranking using the formula .

3. What variables affect a school’s ranking most in a positive way? In a negative way?

4. Use the weights to calculate the US News World & Report score for any college that the user enters the 14 pieces of info for.

**Python Refresher – Be sure to do this before getting started on your project!**



Open Pycharm and start a new project.

Go to File – New and create a new Python file.

Put the csv collegedata file that I’ve sent you in the same directory as this project. You’ll be able to see them on the side:

Always keep these two lines at the top of your file:

**import** numpy **as** np  
**from** scipy **import** linalg

In order to generate a matrix of the college data, type the following:

collegedata=np.genfromtxt(**'collegedata.csv'**,delimiter=**','**, dtype=None)

This command creates a matrix called collegedata that stores the information from the file collegedata.csv. It goes to another column in the matrix whenever it sees a comma between the numbers (that is what delimiter=’,’ does). It also doesn’t declare the table as numbers or letters, since the matrix contains both college names and numbers. (dtype=None). Be sure to use a capital “N.”

You can print the entire matrix by typing:

**print**(collegedata)

You can find how many colleges are in the list by typing:

**print**(len(collegedata))

You can find how many variables are in the list by typing:

**print**(len(collegedata[0]))

You can find out the data for just the Williams school by typing:

**print**(collegedata[0])

How would you find the data for Oberlin?

You’ll probably need to use a for loop in your program, so let’s review those. Remember that for loops require specifying a range. Create a list A=[2,3,5,7,10]. If you wanted to print out one entry at a time, you could type:

A=[2,3,5,7,10]  
**for** i **in** range(len(A)):  
 **print**(A[i])

Remember that range will always stop one short of the maximum value.

Suppose you wanted to find all of the SAT scores in the list, which is contained in the 5th column of the matrix. You could type:

**for** i **in** range(len(collegedata)):  
 **print**(collegedata[i][4])

How would you print out the alumni giving rank?

You might also need if statements. Remember that an if statement needs double equal signs. Suppose you only wanted to print out the even entries in the A=[2,3,5,7,10] list. You could type:

A=[2,3,5,7,10]  
**for** i **in** range(len(A)):  
 **if** A[i] % 2 == 0:  
 **print**(A[i])

You might want to create a 3x2 matrix of zeros. You could type:

B=np.zeros((3,2))  
**print**(B)

Suppose you wanted to create a matrix of zeros but then change a few entries. You could type:

B=np.zeros((3,2))  
B[0,1]=5  
B[1,1]=-4  
B[2,0]=7  
**print**(B)

Suppose you wanted to create a 2x3 matrix of the following entries. You could type:

C=np.matrix([[1,0,4],[3,-1,6]])  
**print**(C)

To find the dimensions of the matrix, you can type:

**print**(C.shape[0],C.shape[1])

Suppose you wanted to create a 3x4 matrix where each Dij entry contained the sum of the i and j indices. You could type:

D=np.zeros((3,4))  
**for** i **in** range(D.shape[0]):  
 **for** j **in** range(D.shape[1]) :  
 D[i][j]=i+j  
**print**(D)

Suppose you have the matrix P=[[1,0,2],[2,0,0],[0,1,2]] and you want to calculate its transpose and inverse. You could type:

P=np.matrix([[1,0,2],[2,0,0],[0,1,2]])  
**print**(P.T)  
**print**(linalg.inv(P))

Suppose you want to multiply P by the vector Q=[[1],[2],[3]]. You could type:

P=np.matrix([[1,0,2],[2,0,0],[0,1,2]])  
Q=np.array([1,2,3])  
**print**(P.dot(Q))

You may also want to use simple lists that are slightly different than numpy arrays.

If you want to create a blank list and then add things to it as you go, you can type:

list=[]

Suppose you want to add the number 7 to the list [10,2,1] and delete the number 2. Then, you want to sort it. You could type:

list=[1,2,10]  
list.append(7)  
list.remove(2)  
list.sort()  
**print**(list)

Suppose you wanted to ask a user for a college name. You could type:

college=raw\_input(**'What college do you want to examine?'**)

Suppose you wanted to ask the user how many colleges they wanted to look at. You could type:

num\_colleges=int(input(**'How many colleges do you want to examine?'**))

**The Three Most Important Things To Remember**

1. I have given you most but probably not all of the commands that will be helpful to you with this project. Google is your friend! If I wanted to calculate the maximum value, for example, in my numpy array, I would google “maximum numpy array” and lots of good information would pop up.

2. Remember to start small and then increase in difficulty level. You might want to work with a smaller file before working with the full college file.

3. Patience is a virtue!! You’ll probably run into many small (and large) problems along the way. Try your best on your own, use Google when you get stuck, and leave your program alone for a while and come back to it with a fresh set of eyes. If all of that still doesn’t help, then ask me for help.